DOCUMENT RESUME

ED 272 549 TM 860 465

AUTHOR Hawley, Willis D.
TITLE Educational Statistics and School Improvement. Statistics and the Federal Role in Education.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE [Oct 85]
NOTE 34p.; In: Invited Papers: Elementary/Secondary Education Data Redesign Project, October 1985; see TM 860 450.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Aptitude Treatment Interaction; *Data Collection; *Educational Facilities Improvement; Educational Improvement; Educational Research; Elementary Secondary Education; Evaluation Criteria; *Information Needs; *Models; *Research Utilization; *School Effectiveness; Statistical Surveys
IDENTIFIERS Effective Schools Research; *Learning Productivity Model; National Center for Education Statistics

ABSTRACT

This paper focuses on how educational statistics might better serve the quest for educational improvement in elementary and secondary schools. A model for conceptualizing the sources and processes of school productivity is presented. The Learning Productivity Model suggests that school outcomes are the consequence of the interaction of five factors: (1) clarity of educational objectives; (2) student characteristics; (3) curriculum technology and instructional strategies; (4) teacher quality; and (5) school environment--opportunity and resources. The importance of educational research is stressed for its ability to identify what types of data are relevant in studying effective schools. The model considers a variety of educational outcomes, both academic achievement and lifetime outcomes or post-school experiences such as occupational success. Two suggestions for data collection are discussed. New data should be used to relate educational processes to educational outcomes, as in the High School and Beyond study. Better coordination of data collection activities is recommended. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
X Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EDUCATIONAL STATISTICS AND SCHOOL IMPROVEMENT

by

Willis D. Hawley

STATISTICS AND THE FEDERAL ROLE IN EDUCATION

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. Nauchamer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED272549

The collection and dissemination of educational statistics have been a federal responsibility for almost 120 years. The federal role in education has changed considerably since then and the demands made for the scope, depth and accuracy of information have increased accordingly. While the appropriateness of the federal government's role in the collection and provision of statistical information is not seriously questioned, debate about what and how data should be collected is recurrent. The reason for this is that statistics are not ends in themselves--they support and facilitate the attainment of other purposes. Many of these purposes are controversial and involve real or potential conflict not only within the federal government but among different levels of government. There seem to be four broad purposes for which educational statistics are used:

1) To provide information on conditions and trends relating to the characteristics and performance of learners and educational institutions,
2) To guide the allocation of resources,
3) To assist in the enforcement of laws and regulations assuring students freedom from racial, ethnic, gender or other invidious discrimination, and
4) To facilitate the improvement of educational policies and practices.

TM 860 465

Until recently, most uses of educational statistics related to the first three of these purposes. However, the current surge of interest in educational reform has seen policymakers and reformers turn to educational statistics to

determine how effective schools have been and what improvements might result in the greatest increases in student learning. And, as more and more people have sought to put educational statistics to work in the cause of change, the limits of available information have become more apparent. As Cooke, Ginsburg and Smith (1985) point out, information available from the National Center for Educational Statistics (NCES) is often inaccurate, incomplete and inconsistent. NCES has no monopoly on such credits. In particular, the extant information tells us too little about the outcomes of education. And, when such information is available, measures that might account for differences among students and school systems often are not.

This paper focuses on how educational statistics might better serve the quest for educational improvement in elementary and secondary schools. In seeking answers to this question, I briefly identify the contributors to the federal information base and suggest that available resources could be better used. I then turn to the notion that if we want statistics to serve the interests of educational improvement, we need to conceptualize the sources and processes of school productivity. Such an analytical model is presented and the types of questions suggested by this framework for which statistical information could provide answers are identified.

Some implications of this approach for current studies undertaken by federal agencies, especially the National Center for Educational Statistics (NCES), and for future inquiries, are explored. I conclude by arguing that the quality of educational statistics can be improved if (a) statistical surveys are theory driven and policy focused and (b) the use of statistics is made easier and more productive.

THE RELATIONSHIP BETWEEN EDUCATIONAL STATISTICS AND EDUCATIONAL RESEARCH

The collection of educational statistics should both be informed by and facilitate research. Statistics, by themselves, seldom tell a very rich story but they can and should encourage better story telling.

I use the term research here to mean the process by which new knowledge that explains why things happen is discovered. To believe that one can understand what is going on in schools by knowing the characteristics of schools, even if one knows the performance of students in those schools, is to be presumptuous or naive--or both.

Statistical gathering and analysis that is uninformed by research encourages simplistic conclusions about the causes of student performance. There are two reasons for this. The first is that important variables may be omitted from the data sets while irrelevant data may be collected. Requests for information that few people will or could use makes unproductive demands on data providers but the omission of variables encourages inappropriate analysis. A second characteristic of statistical information the collection of which is not informed by research is that key variables may be inappropriately specified and measured.

Data that have been shaped by research and collected with theoretical concerns in mind obviously facilitate research. Both the extensive utilization of the National Longitudinal Study of the High School Class of 1972 (NLS-72) and the limitations of these analyses makes this point. In comparison to the NLS-72, the High School and Beyond (HSB) Study was much better informed by research and theory and the potential payoffs from analyses of these data appear to be enormous.

Financial resources available for educational research are meager. At the same time, there is a virtual army of actual and potential educational researchers that could attack "research-influenced" information on education in search of new knowledge. There are ways that the energies of this army could be enlisted and redirected from the ubiquitous mail surveys and meaningless descriptive studies but an exploration of such strategies seems to be the topic of another paper.

Research based on the types of statistical information typically collected by government agencies or their contractees, even data as rich as those collected in the HSB Study, can tell only part of the story. Analysis of large scale statistical data should lead to and be informed by more intensive research that examines the dynamics of student and teacher interaction and otherwise helps us to avoid false assumptions about the meaning of statistical data. For example, the introduction of comprehensive teacher evaluation plans, especially those that involve standardized statewide criteria and processes, is one of the most far-reaching reforms ever introduced in American schools. But it is not the presence of these plans or even the putative characteristics of the plans that will explain any changes in teacher behavior; it is the way these plans are implemented and the data from them are utilized that will tell the tale. One cannot determine how an evaluation system (a curriculum, or other innovations such as a new instructional method, or a management practice) is implemented and how context affects implementation without using research techniques that are more intensive than statistical surveys.

In summary, research and statistical surveys should be seen as having a symbiotic relationship. This is particularly true with respect to statistics that might inform decisions about school improvement.

THE UNDERUTILIZATION OF EXISTING CAPACITY

The National Center for Educational Statistics collects, directly and indirectly, only a portion of the information that might be relevant to educational improvement. Within the Department of Education, a host of other data collection efforts are regularly undertaken. These include various policy studies and program evaluations conducted or supported by the Office of Planning Budget and Evaluation; research and information collected by line agencies such as the Office of Special Education and the Office of Bilingual Education and Minority Language Affairs; the National Center for Research on Vocational Education and the National Institute for Research on the Handicapped; The Office for Civil Rights; and various programs within the National Institute of Education, including the National Assessment of Educational Progress (NAEP). This list is by no means exhaustive. The point is that an enormous amount of information is collected on American education but there is no central effort to plan or coordinate the information collected or even to consolidate it once it is collected. This is not to argue that all data collection efforts within the Department of Education should march to the same drummer. But, the virtual absence of past efforts to define the information needs of the educational improvement mission of the Department of Education denies the Department opportunities to inform the country and influence policy. Various mechanisms have been employed over the years to "control" data gathering efforts but these have been focused primarily on the reduction of paperwork and other burdens imposed on data providers rather than on improving the quality and accessibility of information.

Given the absence of efforts to plan and coordinate data collection and to consolidate statistical data within the department, it is not surprising that the education-related data collection efforts undertaken by other federal

agencies are not part of a comprehensive plan and that data are not assembled in one place. Among the agencies that regularly collect information that might be helpful in understanding and facilitating school improvement are the National Science Foundation, the Department of Labor, the Department of Health and Human Services, the Department of Commerce (especially the Bureau of the Census), and the Department of Defense.

The Secretary of Education could take the initiative in designing a master plan for education-related statistics that would encompass the statistical activities of all of the agencies identified above. The first step in that regard would be to catalog current and planned programs. The second step would be to identify the key variables upon which major studies focus and the uses to which the data are part. A logical outcome of such an initiative would be the establishment of a Federal Interagency Advisory Committee on Educational Statistics. Such a committee could be staffed from the Office of Educational Research and Improvement. An interagency effort could (a) identify sources of data (b) suggest how existing data can be integrated (c) identify areas of unnecessary hindrance and important issues about which data are needed and (d) provide advice to the developers of major new efforts to collect educational information. * One difficulty that will confront such an effort is that the Department of Education is a minor player in the statistics collection game; it does not have a lot to trade with when it seeks cooperation. But the interest in education is great now and, at least for the short term future, the possibility for collaboration may exist.

* NCES has regularly included representatives from the Census Bureau and the Bureau of Labor Statistics on its advisory board and this seems to have been constructive.

States and localities generate much of the data collected by the federal government. In addition, states and localities collect a great deal of information that does not now become part of the federal data base. The growth in the capacity of non-federal education agencies to generate and analyze data appears to have been extensive in the last few years. These "unassembled" data have potentially significant uses in fostering school improvement but this potential is undermined by the absence of standardized definitions of key variables, variation in the information collected from state to state, differences in data collection processes and the difficulty of retrieving the data collected.

A good example of the weaknesses of current state level data is provided by information about teacher supply and demand. At a time when teacher shortages are widely anticipated, states presumably would find it valuable to know how their policies shaping supply and demand compare in effectiveness to those of other states. Some comparative data on state policies are available (cf. Roth and Mastain, 1984). However, data on the consequences of these policies—e.g. number of certified teachers seeking employment, the qualities of these teacher candidates, who actually enters the profession, the volume of reentry, and the rate and character of teacher attrition--are hard to come by even though such data are available for states to secure.

The federal government could take the lead in encouraging and even assisting in the collection of policy-relevant information on the supply and demand for teachers. One mechanism for planning such an initiative is the not overworked Intergovernmental Advisory Council on Education. NCES could be responsible for implementation.

A THEORETICAL FRAMEWORK FOR SPECIFYING STATISTICAL DATA RELATED TO THE IMPROVEMENT OF SCHOOLS

Like data analysis, data collection related to school improvement should be guided by two central questions:

1. What are the problems or issues we want to understand better?
2. What theory or theories might best identify the range of factors that influence the outcomes in which we are interested?

If the answer to the first of these general questions is to know how to most cost-effectively improve schools, the answer to the second question depends on how we would decide a school had improved. I suggest that the central purpose of schools is to produce student learning. Thus, improvement would be measured in terms of amounts or rates of learning or, at least, in terms of the relative achievement of students once all of the variables schools cannot influence are taken into account. I will return below to the types of learning in which we might be interested.

OVERVIEW OF THE MODEL

The effectiveness of most organizations that are in the business of producing something is the consequence of the interaction of five types of factors:

- the clarity and specificity of goals being pursued (e.g. student academic achievement),
- the nature of the raw materials (e.g. the potential for learning students bring to school),
- the sophistication and appropriateness of the technology or means of production (e.g. the curriculum), and
- the quality of the craftmanship (e.g. teacher behavior)
- "environmental" conditions that facilitate or impede productivity

Within this framework, the specific variables to be measured will have to be identified. I will suggest a number of the key factors that current research suggests accounts for variation in student achievement from classroom to classroom and from school to school. This model is based on an extensive recent review of research syntheses and individual studies (Hawley and Rosenholtz, 1984). The relative importance of these factors for enhancing student achievement has been affirmed by a report urging school reform prepared by thirty-two school superintendents from around the country (National Consortium on Educational Excellence, 1984).

There are, of course, other efforts to synthesize current research on school effectiveness (cf. MacKenzie, 1983; Brophy and Good, forthcoming; Purkey and Smith, 1983; Northwest Regional Education Laboratory, n.d.). While there is substantial overlap in the conclusions reached in these reviews, there are also differences. Moreover, the model presented below is derived from research on student achievement measured in conventional terms (i.e., tests of skills and knowledge) and that is a significant limitation. The "Learning Productivity Model" outlined here should be considered illustrative. It seems to be, however, a useful starting point in identifying the types of statistics that will be most helpful in understanding and facilitating school improvement.

Because scholars disagree about the major determinants of school effectiveness and because the importance of different factors depends on the outcome of schooling with which one is concerned, the model or models which would guide federal data collection and coordination efforts should be derived from a consensus of leading researchers and practitioners. The process by which such consensus might be developed is not difficult to imagine: experts should be selected by a snowball reputational technique and they should be provided with alternative assumptions about variables and their interrelationships. The

process should be interactive until agreement is reached either about particular variables or alternative explanations for specified outcomes. Thus, I have chosen not to try to detail specific variables about which one might want statistical data but instead to pose questions which direct attention at the types of data that would be useful in assessing and fostering school improvement.

The Learning Productivity Model of school effectiveness focuses attention on schools and classrooms. This seems an appropriate strategy for organizing one's thinking about school improvement given the substantial evidence that change that affects the quality of education children experience must be generated at the school level if it is to have a significant and continuing influence (Hawley and Rosenholtz, 1984, Ch. 1). Moreover, the frequency with which one sees schools of widely different quality within the same district, even if one takes into account differences in student body characteristics, reinforces the idea that school improvement-related data should provide well textured pictures of schools. However, a number of environmental conditions create the context for school improvement and need to be considered.

The types of factors that influence school productivity and their dominant interrelationships are summarized in Figure 1. Table 1 identifies key variables that comprise each of these factors. Before turning to a discussion of these variables and the directions they point to with respect to the collection and provision of statistical information, the issue of educational outcomes needs to be addressed.

The national interest in school improvement obviously requires that we focus on the ability of schools to produce certain outcomes in the form of student learning. I want to distinguish between these school outcomes from the consequences of formal learning. I will refer to the latter, for lack of a better term as "lifetime outcomes".

SCHOOL OUTCOMES. There appears to be widespread agreement on the importance we should attach to the acquisition of so-called "basic skills" and knowledge about the more or less traditional subject matter measured by tests like the National Assessment of Educational Progress, the Achievement Tests and Advanced Placement Tests administered by the Educational Testing Service, and various standardized tests which report student performance in terms of grade level equivalencies.

There is growing concern, however, reinforced by the poor showing of students on tests of so-called "higher order skills", that our schools are failing to teach students to reason inferentially, to engage in reflection about and systematic analysis of complex problems, and to be creative in the face of uncertainty. Assumptions about the importance of these and similar capabilities to the life chances of individuals and to the economic and social health of the society have yet to be tested. But, there can be little doubt that interest in these types of school outcomes is growing. There can also be little doubt that there is an absence of consensus not only about how to measure these outcomes but how to talk about them with shared meaning. It follows that little is known about how variations in curricula and instructional practices relate to these outcomes. Thus, the federal government could contribute to enlarging the definition of school effectiveness if it could facilitate the development of clearer conceptions of the outcomes by which the effectiveness of schools might be measured. The importance of such an effort, which could take the form of conferences and relatively simple studies of expert and leaders' opinions, is significant because it is possible—as Cooke, Ginsburg and Smith (1985) observe--that the factors that maximize student performance on tests of basic skills may not produce other types of student learning, such as "higher order skills". Indeed, some researchers believe that some elements of the technology of schooling that produces learning of the so-called basics may constrain other learning.

LIFETIME OUTCOMES. Educators, policy-makers and parents have debated the priorities that schools should emphasize since there have been public schools. The nation seems to swing between different priorities with the various advocates invariably making assertions about the effects that one focus or another will have on the experiences and well-being of students once they leave school (cf. Ravitch, 1983). Almost all of this advocacy, however, is unburdened by evidence about the relationship between variations in school experiences and differences in students' post-school outcomes.

While it might be argued that curricula should reflect the values of society, or the community, or the parents independent of the extrinsic consequences they might have for students, this is an unsatisfying position seldom asserted by the advocate of different curricular emphases. Thus, it seems appropriate that the information we have about America's educational systems include data that would allow us to understand how differences in content and levels of investment lead to differences in outcomes beyond academic performance and educational attainment.

Two obvious educational outcomes that it seems useful to know about are occupational success and income. The importance of such information is underlined by a recent synthesis of studies that concludes that individuals' academic performance (as measured by grade point averages or standardized test scores) have almost no relationship to occupational success (Samson, et al., 1984). The central theme of this review of research is reinforced by another recent study showing that a large national sample of employment officers place like emphasis when making hiring decisions place little emphasis on most of the academic outcomes of education that schools are now being asked to focus upon more intently (Crain, 1984).

Other nonacademic outcomes of schooling that might help resolve some of the

recurrent debates about curricula and thus focus reforms on courses of study which benefit students are:

- o participation in the political and social life of communities
- o incidence of antisocial behavior
- o family stability
- o condition of physical and mental health

The importance of considering the goals of schools and the priorities attached to them is critical in understanding the potential for school improvement in any given school or school system. The current educational reform movements, for example, rests on the assumption that the quality of schools has declined. The evidence supporting this assumption is conflicting, however, (cf. Smith, 1984, Hawley, forthcoming) if we focused attention on how well the schools did between 1970 and 1980 in (a) promoting basic skills, (b) meeting the special needs of the disadvantaged and the handicapped, (c) increasing the number of students who did well on Advanced Placement Tests, and (d) reducing the dropout rate and increasing the number of students who attend at least some college, the schools would have to be judged moderately effective. If we ask how well schools did in teaching inferential reasoning and science, or developing students' understanding of democratic process, most schools would have to be judged to have declined in effectiveness. If the explanation for improvements and declines have to do with how schools allocated time and other resources rather than how well they used them, the implications for school improvement would obviously be very different. This rather simple notion that the current dissatisfaction with schools is the result of increased and/or different expectations rather than declines in school quality has received almost no attention in the contemporary public discourse about education in part, perhaps, because there has been little evidence on how school outcomes are related to differences in the priorities given to particular goals and the way

these goals have been pursued.

Recognizing that the factors that affect school productivity may vary depending on the outcome being focused upon, let me identify the general types of influences that current research suggests determine student learning. The statistical information that seems to be both useful and feasible to obtain about each of these influences is indicated by a set of questions that need better answers than we now have. Recall that the Learning Productivity Model has five sets of interactive variables--goals, raw materials, technology, craftsmanship and environmental conditions. Each of these will be considered in turn. I will not cite the literature that links these variables to student learning, unless otherwise noted. The reasons for believing that these factors and others related to the questions posed are important determinants of school outcomes are documented elsewhere (Hawley and Rosenholtz, 1984).

THE GOALS OF SCHOOL IMPROVEMENT

RELATIONSHIPS BETWEEN GOALS AND OUTCOMES. Organizational goals typically focus on products or processes. For our purposes, product goals--the types of learning we want to produce--should be our concern. It might well be interesting to know the fit between process goals and the processes actually being utilized but if we know the latter, which we would if we collected data on the elements of The Learning Productivity Model, we will have this type of statistical information about the processes that are relevant to school improvement.

FIGURE 1

THE LEARNING PRODUCTIVITY MODEL - KEY FACTORS AFFECTING SCHOOL EFFECTIVENESS

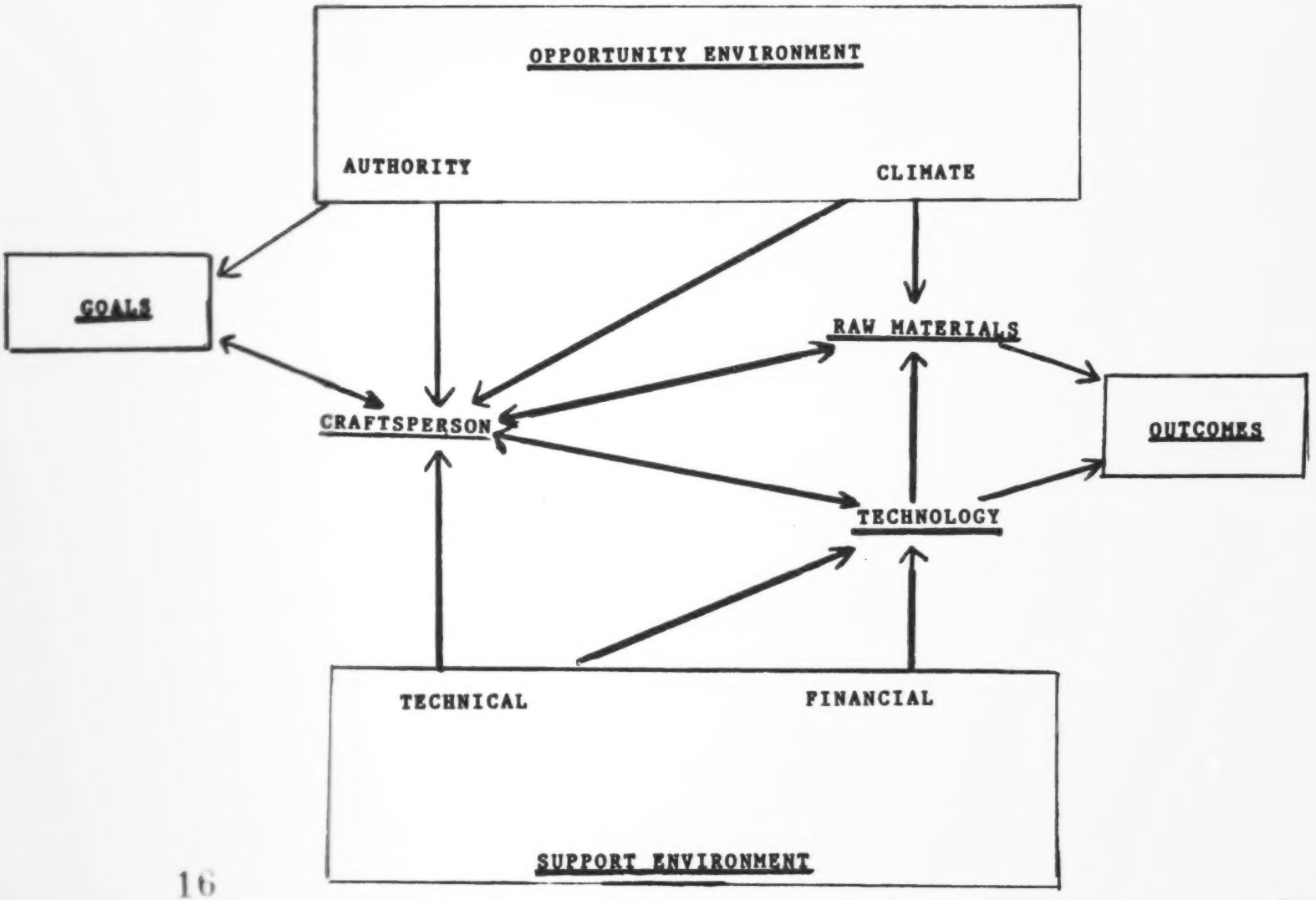

TABLE I

KEY ELEMENTS OF THE LEARNING PRODUCTIVITY MODEL OF SCHOOLING

GOALS

1. CONTENT
2. CLARITY
3. FOCUS
4. CONCENSUS

CHARACTERISTICS AND QUALITY OF
RAW MATERIALS/STUDENTS

1. COGNITIVE CAPABILITIES
2. ACQUIRED KNOWLEDGE
3. MOTIVATION
4. OUT-OF-SCHOOL RESOURCES FOR LEARNING
5. OUT-OF-SCHOOL CONSTRAINTS ON LEARNING

THE TECHNOLOGY FOR PRODUCING LEARNING

1. INSTRUCTIONAL STRATEGIES
2. ORGANIZATIONAL ARRANGEMENTS (EXAMPLES: CLASS SIZE,
3. CURRICULUM
4. TIME
5. LEARNING RESOURCES (EXAMPLES: TEXTS, MATERIALS, TECHNOLOGY, ETC.)

CRAFTSMANSHIP

1. ABILITY AND COMPETENCE
2. CONDITIONS THAT FACILITATE THE USE OF COMPETENCE
3. MOTIVATION

ENVIRONMENTAL CONDITIONS

1. OPPORTUNITIES
2. RESOURCES

One rather straightforward definition of an effective school is that it is a school which achieves its goals. In this sense outcomes and goals can be the same. In order to understand the sources of school effectiveness, however, it will be useful to know the characteristics of the goals to which school personnel subscribe.

CHARACTERISTICS OF GOALS. There are at least four characteristics of the goals to which individuals in a school might profess commitments that appear to be theoretically related to productivity. The first of these is the intensity with which people subscribe to particular priorities. For example, two people may say that reading achievement is their first priority but one may feel that reading supersedes all else by several orders of magnitude. For another, both reading achievement and math achievement may have high priority even though math is ranked lower than reading. Saying that both individuals give their first priority to reading is misleading. Most ways of assessing priorities, however, use simple ranking procedures. (What is needed here is something like the "temperature scales" developed in the voting studies conducted by the Institute for Social Research at the University of Michigan.)

A second dimension of goals that is relevant to student learning is the clarity with which their advocates can describe them. How precise, in other words, are the outcomes being sought? Do seemingly similar goals subscribed to by different teachers have the same meaning for each teacher when the specifics are elaborated?

A third aspect of goals about which data might be gathered is focus. Theodore Sizer may overstate it when he says that the three most important things in explaining school effectiveness are focus, focus and focus. But the

point is well taken (cf, Peters and Waterman, 1983, vis-a-vis the importance of focus in private organizations.) In other words, how many high priority goals do teachers and administrators in a school want to accomplish?

A fourth characteristic of goals which appears to influence school effectiveness is how widely shared the goals are.

THE RAW MATERIAL OF EDUCATION - STUDENTS

Students are the raw material with which schools work. Students bring with them to school different capacities, knowledge and motivation for learning and the effectiveness of schools cannot be measured unless these things are taken into account. Moreover, schools can shape their probabilities of success by influencing the learning readiness of students. Ideally, we would want to know something about the intellectual capacity of students but this is difficult to assess validly under any circumstances and seems beyond the reach of statistical surveys. Other factors that schools might influence or that could be influenced directly by public policies are students' previous knowledge, students' motivation, the resources and assistance students have available out of school, and home or neighborhood constraints on learning. Thus, we might better understand the effectiveness of schools and the promise of different improvement strategies if we had data that spoke to the following questions.

What was the performance of the students in their previous school or class with respect to the outcomes in which we are interested? What is the socioeconomic status of the students? What types of preschool learning experiences, if any, did students have? What types of learning-relevant interactions do students have with their parents, older siblings and other adults with whom they spend considerable time? What proportion of the students are from single parent families and how many have a parent at home when school

is out? What does the school and teachers do to involve parents in the education of their own children? How much and what kind of homework do students do and under what conditions can they study at home? What kind of supports for and constraints on learning are there in the students' immediate neighborhood? How much time do students spend watching television and what do they watch?

THE TECHNOLOGY OF EDUCATION

When students go to school they experience a range of structures, processes and learning resources. These "means of production" --which obviously vary considerably among states, school districts, schools and classrooms--make up the technology of schools. Of course, the effectiveness of the technology in producing student learning is significantly influenced by how well it is used but we can make policy—and do—about what I am calling technology more or less independently of the policies we make that affect the quality and performance of teachers and administrators.

Data collected by NCES and other federal agencies are typically short on information about what actually happens in schools. The HSB data go a long way in addressing this limitation of available statistical evidence but, of course, they tell us only about high schools at two points in time. The types of information about school technology that research suggests would be helpful in assessing school productivity and developing improvement strategies include data on instructional strategies, organization arrangements, curriculum, the use of time, and learning resources. Some of the questions we would want this information to address are set out below.

INSTRUCTIONAL STRATEGIES. What mix of strategies--such as whole class teaching, programmed instruction, individualized learning, cooperative learning, mastery learning, peer tutoring and interactive teaching--is employed and how

does this vary by subject matter and learning objectives? How often are students evaluated and how are decisions made about movement through the curriculum and grade levels? How high and how clear are standards of student performance and what types of rewards do students receive for meeting standards?

ORGANIZATIONAL ARRANGEMENTS. What is the class size with and without teacher aides? How many students are there in the school and how are they assigned to instructional units (e.g. grades or "houses" or blocks)? What grouping and tracking procedures are employed? What is "skewness" of student performance in particular classrooms?

CURRICULUM. What subjects are taught, at what level of difficulty and for how many hours each day or week. Does the school have a core curriculum that is well articulated across grade levels? Are the tests used to measure student performance articulated with the curriculum?

TIME. How many hours of actual instruction are allocated each week free from interruption? How many hours do teachers teach each week? How many days each year do students attend school?

LEARNING RESOURCES. What is the condition and nature of the instructional facilities? What support systems do teachers have—library, volunteers, audio-visual, business-school linkages, etc.--and how often are they available and used? How many computers are available and how are they used?

THE QUALITY OF THE CRAFTSPERSON

A decade or so ago, researchers and policymakers--but not parents--asked questions like "do teachers make a difference"? The recent outpouring of research on school effectiveness leaves little doubt that both teachers and

administrators make a big difference in the quality of education students receive. (cf. Hawley and Rosenholtz, 1984, Chs. 1-3). For example, not only do teachers influence how students learn by the way they implement a curriculum, teachers also shape what students are taught (Brophy, 1980; Green and Harker, 1982). Recent research also suggests that teachers significantly influence parental involvement in their children's education (Epstein, 1984).

Statistical data about educators is very limited. It is not possible, for example, to get a clear idea of what the career paths of principals and administrators look like under different circumstances. Little is known about the qualities of those who actually teach or what difference these qualities make in student performance (cf. Evertson, Hawley and Zlotnik, 1985). Very little is known about teacher attrition. And so on. (The current activity within NCES to develop a better picture of teachers is ambitious and well conceived and it will be much welcomed by policymakers. A similar effort with respect to school administrators is also needed.)

There are three interrelated aspects of craftsmanship which appear to influence student learning: ability and competence, conditions that facilitate or impede the use of competence, and motivation. Some questions to which it seems important to have answers about these interrelated aspects craftsmanship follow.

ABILITY AND COMPETENCE. What are the levels of general academic ability and subject matter knowledge that teachers and principals possess? How much of what types of formal training have teachers and principals had and at what stage of their career did this training occur? How many and what kind of opportunities for increasing their professional competence have teachers and

principals had? What opportunities have teachers and principals had to learn informally from their peers?

MOTIVATION. What priorities do teachers say they give to the different reasons why they teach? What is the volume and character of feedback about their performances that teachers and administrators receive? Is there a formal evaluation system? What are its criteria and how is it used? What is the salary range that is accessible and on what bases are salary increases awarded? Are there financial incentives available other than salary increases? Do teachers and administrators have an opportunity to participate in key decision affecting professional practice? In what ways is superior professional performance recognized and rewarded (besides economic rewards)?

CONDITIONS THAT FACILITATE EFFECTIVE PRACTICE. What level of support do teachers and administrators receive from their respective administrative superiors? How often do teachers have the opportunity to interact professionally with their peers and do norms of collegiality exist? What is the level of order and discipline in the school? Is the teaching time of teachers protected from interruptions and diversion? How widely shared are goals for student learning and at what levels of expectation are these goals pitched? What are the number of students in the school?

It should be noted that NCES has focused increasing attention on teachers in recent months and several surveys have sought to collect data about teachers and what they do. While this effort addresses several of the questions raised above, the bits and pieces apparently cannot be aggregated and too little attention is focused on outcomes. The HSB data do provide student outcome data but little information about teachers' careers and personal characteristics. The NCES labor market survey provides information about teacher shortages and teacher incentive plans but no information about salaries.

ENVIRONMENTAL CONDITIONS

The problem of controlling for non-school factors that affect the productivity of a school cannot be dealt with adequately except under experimental conditions. Therefore, the theory that guides data collection must encompass critical student characteristics beyond the reach of the school and aspects of the schools' ecology. Much has been written about the factors external to individual schools that influence school improvement. Unfortunately, aside from a handful of case studies, little emperical research on the relative importance of these factors exists. Thus, unlike other dimensions of the Learning Productivity Model, this one is not researched-based.

The collection of data relating to school external environments is problematic not only because so little evidence is available upon which to select from among the theoretically interesting ecological factors that could influence student learning but because of the difficulty of acquiring the information. If we were to constrain the collection of data to those that could be provided by occasional national samples or by school systems from existing information bases, the types of information relevant to school improvement that might be collected seems a bit easier to think about.

It seems reasonable to think of environmental conditions that influence school productivity as being of two broad types: (a) conditions that grant or constrain the opportunities of educators and students and (b) resources available with which educators and students can pursue their goals.

THE OPPORTUNITY ENVIRONMENT. The opportunity environment essentially shapes ambitions and expectations and grants--formally or informally—authority to act. Its' dimensions and their consequences might be understood if questions

such as the following could be answered. How much decision-making discretion do school administrators and teachers have to make and implement policy? How much support do the goals of school-level personnel have among district-level administrators and other teachers? How stable are residential patterns and pupil assignments to schools? How available to students are low-cost options to pursue postsecondary education? What types of of employment opportunities in the community can students expect to have upon graduation?

THE RESOURCE ENVIRONMENT. The resources available to pursue the ambitions and expectations held by individuals and by the collective professional and student populations of school should affect student learning. Resources, of course, create opportunities but the "opportunity environment" just noted relates to sources of motivation provided by the environment while the "resource environment" provides capabilities that facilitate action relevant to the individual and organizational purposes that have been discussed as deriving from factors that characterize the school and its student body.

The nature of the resource environment might be understood if answers were available to several questions. What is the level of expenditures per pupil taking into account the number of students with needs for resource-intensive programs? What is the quality and quantity of technical assistance available in the district to implement school-level goals? How many individuals participate in school activities as volunteers in support of instructional or extracurricular activities? What resources, if any, are provided to schools by business and public and nonprofit agencies? What is the "tax-effort" (the extent to which the available tax base is utilized) of the state and the school district? This last indicator may be a surrogate measure of the school environment's commitment to public education.

SUMMARY

I have argued that the collection of statistical data related to school improvement should be guided by grounded theories and hypotheses that might explain differences in student learning among schools and classrooms. A conceptual framework implicitly embodying such explanations, which I have called the Learning Productivity Model, has been outlined. Identifying the key elements of models and questions that might profitably guide data collection and analyses does not, of course, add up to specific recommendations about priorities that should be placed on gaining particular information that would facilitate the development of effective school improvement strategies. And, obviously, the ways such variables might be measured are only hinted at in the discussion above. My presumption, however, is that these next steps in the development of a plan for collecting statistical information is given direction by this model building exercise.

CONCLUSION - NEXT STEPS

Let me conclude by briefly discussing two courses of action which would result in statistical data bases that would support the formulation of productive school improvement policies and practices: (a) the development of a plan for collecting new types of data (b) the integration and enrichment of existing statistical information.

THE COLLECTION OF NEW DATA

In general, statistical data related to education that is now available is not particularly helpful in developing new policy. The current array of educational statistics focuses primarily on the condition of education rather than on explanations for that condition. In other words, little effort has been

made to relate educational processes to educational outcomes. Moreover, as noted above, when outcomes are assessed, the range of educational outcomes measured is narrow. For example, the 1985 Teacher Questionnaire, while it deals with how teachers spend their time, barely touches on the teacher characteristics, behavior, and circumstances most powerfully related to student learning and provides no outcome data whatsoever about either students or teachers. Of what value is such information? If it does lead somewhere, the probability that it will lead us down primrose paths is at least as great as the prospect that the road it will direct us to will be paved with yellow brick.

This generalization about the atheoretical character of data collection does not apply to all surveys. The most notable exception is the HSB Study. There are lessons that this study has for new data collection efforts relating to school improvement even though that is not the primary purpose of the study. First, schools should be seen as social systems and both formal and informal processes should be assessed. Second, an expert review panel that guides the design and redesign of statistical studies seems to be cost-effective. Thus, the next step would be to convene such a panel, refine the theoretical framework, specify variables, consider the sampling issues and otherwise develop the scope of the project so that the potential benefits and the costs can be estimated. An HSB-type study will be very expensive but the return on investment, in terms of usable knowledge and educational improvement will probably be high, if early evidence on the use of the HSB Study is any indication. Comprehensive data collection efforts like the HSB Study or a similar one which would follow from the Learning Productivity Model would be much more valuable to policy makers and would-be reformers if data were collected that would allow analysts to understand the economic costs of alternative improvement strategies.

Because the statistical error that threatens all data analyses is reduced

to the extent that one can take into account all of the alternative explanations for a given phenomenon, and because we typically want large data sets to serve multiple purposes, an obvious problem faced by those who design statistical studies is knowing what data NOT to collect. The potential cost/benefit ratio of data collection itself is one way to make such decisions and theory as well as long run utilization studies can be used to make such calculations. The probable reliability of the data is another factor that should be considered and it would seem useful to develop, perhaps through expert interviews, an understanding of the sources of error in survey responses that could be used across studies. For example, such considerations might include the probability that individuals would perceive themselves or a goal they value being advanced (or the converse) by accurate provision of data and the degree to which the information is readily available.

THE INTEGRATION AND ENHANCEMENT OF EXISTING DATA

I have argued that most statistical surveys provide only part of the picture we need to have of how schools operate and the outcomes they produce. This proposition holds for statistics that address issues other than school improvement. There are good reasons for this having to do with cost, burdens on providers, etc., but among the several studies conducted or sponsored by the federal government, there are many pieces of the puzzle. Moreover, several studies could be enriched if data available from published sources other than statistical studies were added to various data sets. NCES, therefore, might consider the following strategies for making available statistics and recurrently conducted studies more accessible and more usable:

(a) coordinate the data gathering of agencies within the Department of Education, across the federal government, and between federal agencies and other levels of government.

(b) "nest" future studies in such a way that data from the same sites could be integrated. For example, it might be possible to conduct the HSB and National Assessment studies in the same or overlapping locations.

(c) identify common sites at which data has already been collected. For example, the teacher supply and demand study may have been conducted in locations for which data on teacher salaries were collected.

(d) enrich existing statistical data by adding information from other sources. For example, the teacher demand and shortage study could be enhanced by adding information on career entry requirements (available from The National Association of State Directors of Teacher Educational Certification).

Steps such as these, coupled with efforts to prepare the data in formats that are well documented and readily usable by researchers and practitioners, would increase the frequency and sophistication with which educational statistics were analyzed. The more analysis that is carried out, the more will be known about the strengths and weaknesses of the data collected. Such knowledge would result in improvements in the quality of the data collected in future studies.

SO, WHO WILL FUND ALL OF THIS?

What I have called for in this paper could easily be dismissed as being beyond the resources of the Department of Education and therefore unfeasible. Surely, anyone who suggests more federal activity these days would seem to be out of touch. There are two general ways to respond to the legitimate concern

about new expenditures. The first of these is to do more by making better use of current resources. The second is to convince policymakers that spending money on statistics will not only improve learning but will facilitate more cost effective policy decisions. Let me consider the second of these arguments first.

It is worth noting that the Office of Management and Budget (OMB) often "discourages" statistics gathering proposals on the grounds that they lack policy relevance. What is being proposed here is to increase the relevance of statistics to policy making.

All levels of government expend about $120 billion on public elementary and secondary education. The NCES budget is an infinitesimal proportion of this and if the costs of all other education data gathering efforts were added up, the relative level of investment in statistics would still be minuscule. If the information developed from the types of school improvement-related data I have urged be collected were to better inform the decisions of one large school system, the cost of the national effort might be covered. Among the types of decisions that might yield large dollar returns that could be informed by the sophisticated analysis of quality data of the sort discussed above are choices about class size, beginning teacher salaries, the use of economic incentives, investments in traditional staff development and other matters. Of course, analyses could point us toward higher expenditures but, if so this would result in greater productivity. If investments in schooling could be tied to post-school outcomes, the potential for return on the investment in statistics would be very great. Of course, statistical data seldom provide, by themselves, clear policy directions. But they can challenge myths, they can suggest new options, they can confirm predispositions and they can identify issues worth further investigation. Moreover, when cynics charge that statistics can be misleading

and often provide incomplete pictures the answer might be: compared to what? Most policy choices are made on the basis of intuitions informed by predispositions and by the concern decisionmakers for their political and occupational futures. Good statistics can serve as a balance wheel, if not a steering wheel, in the policy process.

What about making do with the resources we have? To pursue the course suggested in the heart of this paper, it would be desirable, perhaps necessary, to conduct one study a year of the size and sophistication of a "wave" of the HSB Study. Multi-year longitudinal research does not seem necessary though, of course, that too would be valuable. One possible source of those funds is to discontinue work being done now that does not seem to go anywhere. In addition, it might be feasible to combine resources from several of the NIE centers whose missions overlap the purposes to which school improvement-relevant statistics could be put.

It might also be possible, especially in view of the recent position of the chief state school officers to use the NAEP data comparatively, to connect the study of school processes in some way to the outcomes being studied in the NAEP. This would probably require in a larger sample and increased costs to conduct the NAEP but such piggy-backing would reduce the overall expense of the school improvement study even if this study bore the increased costs of the NAEP.

If a major new study or set of studies was not possible, two other options remain. One is for NCES to take the technical lead and provide some financial

incentive that would encourage separate state studies that would be conducted, in part, in accord with a common design and with common basic variables being measured. A second option would be to do those several things noted earlier in the conclusion related to making more effective use of existing data.

Where there's a will, . . .

References

Brophy, J.E. (1980). Teachers' cognitive activities and overt behaviors (Occasional Paper No. 39). E. Lansing: Michigan State University, Institute for Research on Teaching.

Brophy, J.E., & Good, T.L. (1974). Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart & Wilson.

Brophy, J.E., Monias, J., & Good, T.L. (in press). "Teacher Behavior and Student Achievement," in M. Wittrock (Ed.). Handbook of Research on Teaching (3rd ed.). New York: MacMillan.

Cooke, C., Ginsburg, A., & Smith, M. (1985). The sorry state of educational statistics. Paper prepared for the U.S. Department of Education, January.

Crain, Robert, L. (1984). The quality of american high school graduates: What personnel officers say and do about it. Center for the Social Organization of Schools Report. Baltimore: The Johns Hopkins University.

Epstein, J.L. (1984). Single parents and the schools: The effect of marital status on parent and teacher evaluations (Report No. 353). Baltimore: The Johns Hopkins University, Center for Social Organization of Schools.

Evertson, C., Willis, D., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. Journal of Teacher Education, 36 (May - June) forthcoming.

Green, J.L., & Harker, J.O. (1982). Gaining access to learning: Conversational, social, and cognitive demands of group participation. In L. Cherry-Wilkinson (Ed.). Communicating in the classroom. New York: Academic Press.

Hawley, W.D., (Forthcoming). False premises, false promises: The mythological character of public discourse about education. Phi Delta Kappan."

Hawley, W.D., & Susan Rosenholtz (1984). "Good schools: what research says about improving student achievement", Peabody Journal of Education, 61, entire issue.

MacKenzie, D.E. (1983). Research for school improvement: An appraisal of some recent trends. Educational Researcher, 12 (4), 5-17.

National Consortium on Educational Excellence (1984). An Agenda for Educational Renewal: A Fiew from the Firing Line. Nashville: Peabody College of Vanderbilt University.

Northwest Regional Educational Laboratory (n.d.). Effective School Practices: A Research Synthesis. Portland, Oregon.

Peters, T.J., & Waterman, R.H., Jr. (1982). In Search of Excellence: Lessons from America's Best-run Companies. New York: Harper & Row.

Purkey, S.C., & Smith, M.S. (1983). Effective schools—a review. Elementary School Journal, 16, 27-29.

Ravitch, D. (1983). The Troubled Crusade. New York: Basil Books.

Roth, R., & Mastain, R. (Eds.) (1984). Manual on certification. Sacramento: National Association of State Directors of Teacher Education and Certification.

Samson, G.E., Grave, M.E., Weinstein, T., & Walberg, H.J. (1984). Academic and occupational performance: A quantitative synthesis. American Educational Research Journal, 21, 311-322.

Smith, Marshall S. (1984). Educational Improvements Which Make A Difference: Thoughts About the Recent National Reports About Education. Washington, D.C.: Federation of Behavioral, Psychological and Cognitive Sciences.